APPENDIX B

Data Analysis Approach

APPENDIX B

Data Analysis Approach

The approaches used to analyze the data from measurements of elements in used motor oil and properties of the used oil are summarized in this appendix:

- Modeling wear rate of engine metals
- Statistical analysis of engine metal removal rates
- Statistical modeling of oil properties.

Modeling of Engine Metals

The amount of metals removed from the engine during each oil change interval is calculated from the reported concentrations of engine metals in the used oil and information on oil consumption determined from oil additions and dipstick levels at each oil change. The calculations are based on differential equations derived from the following assumptions:

- Elements are entering the engine oil at a constant rate
- Oil is leaving the engine (through leakage or combustion exhaust) at a constant rate
- As oil is lost from the engine, the elements contained in the oil are lost at the same rate
- Oil is added in one-quart increments, and it is not added until the oil level is down one quart.

In graphical form, the assumed behavior of the oil concentration level is shown in Figure B-1. This example graph represents the predicted element concentration history for a vehicle that was driven 3,000 miles, had a quart of oil added at 2,000 miles, had an initial element concentration of 0 ppm, and a final concentration of about 17 ppm. Notice that although the rate of material entering the oil is assumed constant, the concentration increase is not linear with time. This effect is due to the decreasing volume of oil in the sump. Because the concentration of critical engine metals in the oil is known only at the beginning of an interval (when fresh oil is added) and at the oil change (through the spectrochemical oil analysis), a way of calculating the concentration at any time during the interval was needed to predict the amount of metals leaving the engine. The differential equation that describes this behavior and important steps in its derivation are shown below.

By definition

$$C = \frac{M}{V} \tag{B-1}$$

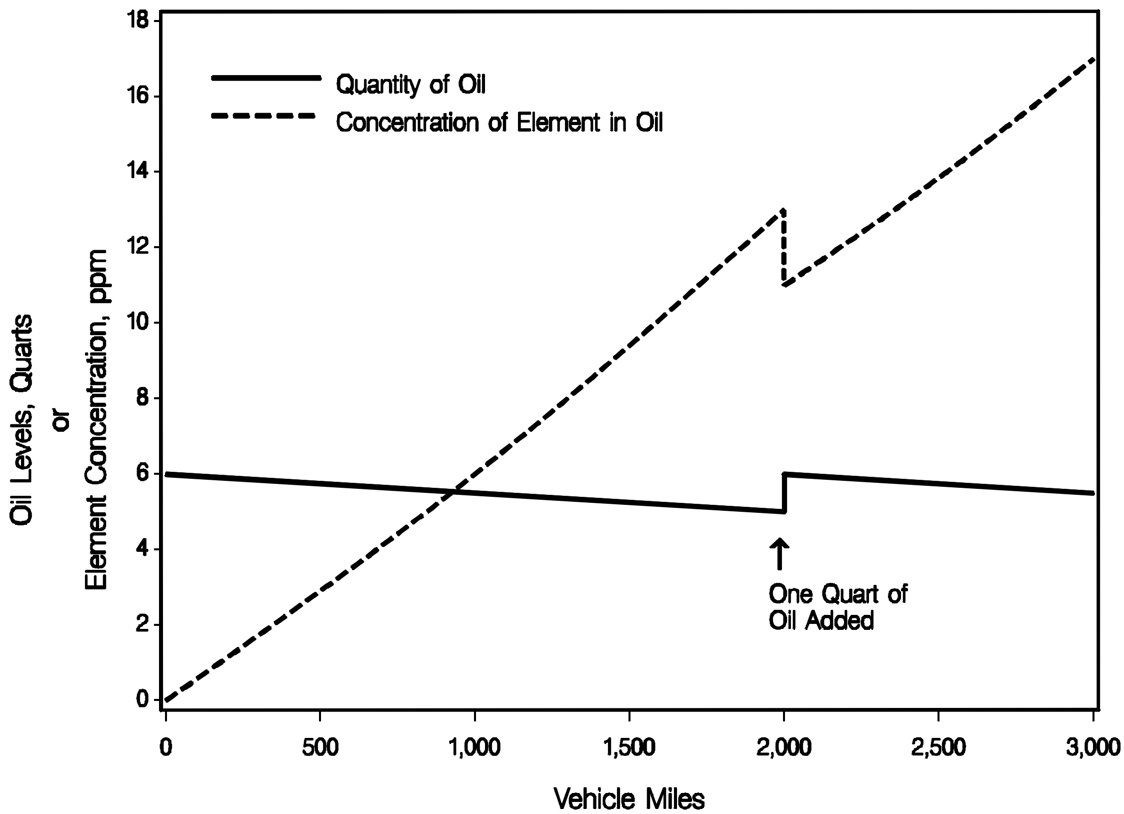

Figure B-1. Behavior of Element Concentrations in Engine Oil

where

C = concentration by volume of the element in the oil
M = mass of the element in the oil
V = total sump volume.

Differentiating one obtains

$$dC = \frac{VdM - MdV}{V^2} \tag{B-2}$$

Also by definition

$$dV = -Q_O dt \tag{B-3}$$

and

$$dM = -Q_O C dt + Q_E dt \tag{B-4}$$

where

Q_O = the volume flow rate of oil leaving the engine
Q_E = mass flow rate of the element into the oil.

Substituting equations (B-3) and (B-4) into equation (B-2) and simplifying one obtains

$$dC = \frac{Q_E\,dt}{V} \tag{B-5}$$

and by integration

$$C = C_O - \frac{Q_E}{Q_O} \ln\left(1 - \frac{Q_O}{V} t\right) \tag{B-6}$$

This is the basic equation for any interval between oil additions or changes. For example, in Figure B-1, equation (B-6) defines the concentration at the 2,000 mile mark. Between 2,000 miles and the oil change at 3,000 miles, the same equation can be used except that the initial concentration, C_O, is equal to the final concentration of the first interval, C, multiplied by a dilution factor for the added quart of oil. Solving for Q_O and summing the various intervals between oil additions, the general solution is:

$$Q_E = \frac{Q_O\left[C_O - \left(\frac{V}{V-1}\right)^N C_E\right]}{\left[\ln\left(1 - \frac{Q_O}{V}\tau\right)\right]\sum_{i=0}^{N}\left(\frac{V}{V-1}\right)^i + \left(\frac{V}{V-1}\right)^N \ln\left(1 - \frac{Q_O}{V}\tau_r\right)} \tag{B-7}$$

where

N = number of one-quart oil additions
τ = number of miles between one-quart oil additions
τ_r = number of miles between the last one-quart oil addition and the oil change.

In this analysis, the concentration of elements is in terms of mass of element per mass of oil (rather than volume of oil), so the volume of oil must be changed to a mass of oil by including a density term ρ as shown below.

$$Q_E = \frac{Q_O\,\rho\left[C_O - \left(\frac{V}{V-1}\right)^N C_E\right]}{\left[\ln\left(1 - \frac{Q_O}{V}\tau\right)\right]\sum_{i=0}^{N}\left(\frac{V}{V-1}\right)^i + \left(\frac{V}{V-1}\right)^N \ln\left(1 - \frac{Q_O}{V}\tau_r\right)} \qquad \text{(B-8)}$$

In order to calculate the terms, τ, τ_r, and Q_O, the following equations are given.

$$\tau = \frac{M_F - M_I}{N + D} \qquad \text{(B-9)}$$

$$\tau_r = M_F - M_I - (N\tau) \qquad \text{(B-10)}$$

$$Q_O = \frac{N+D}{M_F - M_I} \qquad \text{(B-11)}$$

where

M_F = final mileage in the oil change interval
M_I = initial mileage in the oil change interval.

The term, D, represents the amount of oil below a full sump present at the oil change. For example, D would be 0.5 for a vehicle that is one-half quart low at the oil change.

The above equations enable one to determine an average, constant removal rate of material from the engine over each individual oil change interval. By observing the removal rate of critical engine elements for each oil change, trends for normal behavior for the engine quickly become established.

Statistical Analysis of Engine Metal Removal Rates

Statistical comparisons of the removal rates of metals were made between each alternative fuel fleet and its control fleet. (Control fleets consist of vehicles from the same manufacturer but operating on regular unleaded gasoline.) Comparisons were based on the average weight of metal removed from the engines at various mileage levels. However, prior to performing this comparison, a number of preliminary analyses were needed. Because of the statistical advantages of combining data from control vehicles at different demonstration sites, statistical regression analysis was performed to determine if the rates of engine metal removal among control vehicles from the same manufacturer are consistent across sites. Such differences could occur as a result of differences in duty cycles or maintenance practices at the demonstration sites. The analysis did not reveal any significant site-to-site differences.

Out of the 918 oil changes reported, 60 oil changes were performed in which oil samples were not collected. Because many of these missing data were from the initial oil changes, it was necessary to estimate values for vehicles with missing data using the estimated metal removal rates from vehicles in the same fleet with complete data. This was done by fitting regression models to the available data from vehicles in the same fleet. Then, the estimated metal loss rate was used to “impute” values for vehicles with missing data. The regression model was

$$TM_{ijk}\ (\text{grams}) = R_{ij} \times OM_{ijk} + e, \tag{B-12}$$

where

TM_{ijk} = total engine metal removed during the jth oil change interval from vehicle k of fleet i

R_{ij} = rate of engine metal removal per mile during the jth oil change interval for fleet i

OM_{ijk} = number of miles driven between the j-1st and jth oil change for vehicle k of fleet i (OM is referred to as "oil miles")

e = random effect associated with measurement error and differences among vehicles within the same fleet.

The rationale for this approach is that it permits the maximum use of available data. Without this approach, the data from oil changes following the one with missing data could not be used to calculate cumulative metal removal. Because the level of data completeness achieved was 93 percent, the potential bias from this method is expected to be minimal.

The next step in the analysis was to calculate for each vehicle the cumulative weight of each metal removed from the engine at each oil change. For each vehicle the cumulative weight removed was calculated by

$$CWM_{ijk} = \sum_{l \leq j} TM_{ilk} \tag{B-13}$$

and

$$COM_{ijk} = \sum_{l \leq j} OM_{ilk} \tag{B-14}$$

where

CWM_{ijk} = cumulative weight of metals removed from vehicle k of fleet i by the jth oil change

COM_{ijk} = total mileage on vehicle k of fleet i at the jth oil change (cumulative oil miles).

Plots of CWM versus total miles (COM) were reviewed to determine the level of consistency among vehicles of the same fleet and to identify statistical outliers. The plots showed a high degree of consistency among vehicles.

Finally, for each fleet, the average cumulative metal removed at specific mileage levels (2,500, 5,000, 10,000, 15,000, and 20,000) was calculated using interpolated values from individual vehicles. Regression methods were used to estimate average levels for each fleet, determine the precision of these averages, and identify significant differences in the averages between each alternative fuel fleet and its control fleet.

The degree to which the number of miles between oil changes affects metal accumulation rates was also investigated. Plots of the estimated cumulative weights of selected metals versus the average miles between oil changes were prepared for each fleet. This was done because there were substantial differences in the number of miles between oil changes among the various fleets, including fleets of control vehicles at different demonstration sites. The presence of a trend within fleets would indicate that the length of the oil change interval should be considered in the statistical treatment of the data. No such trend was observed.

Statistical Modeling of Oil Properties

The oil properties monitored at each oil change include total base number, viscosity, nitration, and oxidation. The relationship between these properties and the miles driven since the last oil change is of particular interest to fleet operators because of the potential impact on the preventive maintenance schedule.

For each fleet, the values of each property were initially plotted against oil miles (miles since the last oil change) to determine if there were significant trends. The potential effects of cumulative vehicle miles on the properties was also investigated. Several empirical models were fitted to the data and tested for goodness-of-fit. The best-fitting models for total base number (TBN), viscosity (V), and nitration (N) at each oil change are

$$\log(TBN/TBN_O) = \beta_1 \text{ x } OM + \epsilon_1, \tag{B-15}$$

$$\log(V) = \alpha_2 + \beta_2 \text{ x } OM + \epsilon_2, \tag{B-16}$$

and

$$N = \alpha_3 + \beta_3 \text{ x } OM + \epsilon_3, \tag{B-17}$$

where TBN_o is the average measured TBN in unused oil samples; OM is the miles driven since the last oil change; $ß_1$, α_2, $ß_2$, α_3, and $ß_4$ are constants; and ϵ_i (i= 1,2,3) are independent random errors that are assumed to be approximately normally distributed. Separate models were fitted to the measured properties at the initial oil change, if warranted.

Statistical regression analysis was used to test for significant trends (i.e., nonzero slope ß) for each fleet and to compare the oil properties for each alternative fuel fleet with the corresponding control fleet. The comparisons were based on the predicted oil properties at 3,000 miles following an oil change.